

# Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases *in vitro* and *in vivo*

Rodolphe Suspène\*, Denise Guétard\*, Michel Henry\*, Peter Sommer†, Simon Wain-Hobson\*\*, and Jean-Pierre Vartanian\*

\*Molecular Retrovirology (Centre National de la Recherche Scientifique Unité de Recherche Associée 1930) and †Cell Biology of the Nucleus Units, Pasteur Institute, 28 Rue du Dr. Roux, 75724 Paris Cedex 15, France

Edited by Francis V. Chisari, The Scripps Research Institute, La Jolla, CA, and approved April 11, 2005 (received for review November 5, 2004)

Because the replication of hepatitis B virus (HBV) proceeds via an obligatory reverse transcription step in the viral capsid, cDNA is potentially vulnerable to editing by cytidine deaminases of the APOBEC3 family. To date only two edited HBV genomes, referred to as G → A hypermutants, have been described *in vivo*. Recent work suggested that HBV replication was indeed restricted by APOBEC3G but by a mechanism other than editing. The issue of restriction has been explored by using a sensitive PCR method allowing differential amplification of AT-rich DNA. G → A hypermutated HBV genomes were recovered from transfection experiments involving APOBEC3B, -3C, -3F, and -3G indicating that all four enzymes were able to extensively deaminate cytidine residues in minus-strand DNA. Unexpectedly, three of the four enzymes (APOBEC3B, -3F, and -3G) deaminated HBV plus-strand DNA as well. From the serum of two of four patients with high viremia, G → A hypermutated genomes were recovered at a frequency of  $\approx 10^{-4}$ , indicating that they are, albeit relatively rare, part of the natural cycle of HBV infection. These findings suggest that human APOBEC3 enzymes can impact HBV replication via cytidine deamination.

hypermutation | DNA editing

G → A hypermutated retroviral genomes result from the editing of nascent DNA by APOBEC3 cytidine deaminases (1–7). This was originally demonstrated for HIV-1 and the APOBEC3G member of a cluster of seven genes (3A–3H) on human chromosome 22 (8, 9). Editing occurs on the background of a  $\Delta$ vif genome (1–5, 10). The HIV Vif protein prevents packaging of either APOBEC3F or -3G resulting in their ubiquitination and degradation by the proteasomal pathway (11–14). Editing of cytidine residues in neo-synthesized minus-strand cDNA results in the formation of multiple uracil residues that are read as T during plus-strand synthesis. Albeit referred to as G → A hypermutants by reference to the viral plus strand, mechanistically the action is occurring on the minus strand and independently of reverse transcriptase (7). Up to 60% of G residues in a lentiviral genome can be substituted resulting in the total loss of information (15–17). As such, APOBEC3F and -3G constitute a powerful restriction mechanism to reverse transcription. Retroviruses have either to avoid replication in cells expressing APOBEC3 molecules or else evolve a mechanism that neutralizes their effect. The *vif* gene of human and primate lentiviruses, as well as their homologues in most of the other lentiviruses, reflect the latter solution.

Hepatitis B viruses (HBVs) replicate via an obligate reverse transcription step occurring in a capsid structure close to the endoplasmic reticulum. A pair of G → A hypermutated genomes were identified in the blood of a chronically infected patient yet have remained unique despite a burgeoning database (18). Furthermore, gene chip analyses of liver tissue have failed to show significant expression of APOBEC3 molecules, apart from APOBEC3C (ref. 19; <http://genecards.bcgsc.bc.ca>). Together, these findings suggest that HBV might have adopted the alter-

native route and sought replication in cells with little or no APOBEC3 expression.

Recent reports demonstrated that HBV replication could be strongly restricted by APOBEC3G and -3F in an experimental setting (20, 21). Restriction was highlighted by a strong reduction in the proportion of pregenomic RNA and DNA in the cytoplasm. Despite this, no G → A hypermutated genomes were identified, suggesting that the two enzymes curtailed replication via a nonediting mechanism. Using the same experimental system, another group confirmed restricted HBV replication but, like the original report, could not find a significant increase in the proportion of G → A hypermutants when the hepatoma Huh7 cell line was used (22). However, using the HepG2 cell line they found a 5-fold increase in the frequency of G → A hypermutants. Yet it is not obvious that there was a cell line effect. When Huh7 was cotransfected by an HIV-1 $\Delta$ vif genome and an APOBEC3G expression plasmid, restricted HIV-1 replication was observed when the supernatant was cultured on a susceptible cell line (20, 21).

These findings suggested that HBV replication may indeed be restricted in an experimental context by APOBEC3G, perhaps in an unconventional manner. However, the discrepancies between the findings and the dearth of such G → A hypermutated genomes *in vivo* suggests that the picture is incomplete.

Recently, a novel PCR technique was reported that allows differential DNA amplification of G → A hypermutants (23). This technique relies on the fact that the DNA of an AT-rich variant will melt at a slightly lower temperature than parental DNA. Using this technique, we have investigated the effect of various APOBEC3 family members on the integrity of neo-synthesized HBV DNA in Huh7 cells. The results unequivocally demonstrate that HBV DNA is sensitive to extensive cytidine deamination by four APOBEC3 molecules. When the technique was applied to DNA extracted from the serum of four chronically HBV-infected patients, numerous G → A hypermutants were identified in two cases. In short, HBV G → A hypermutation parallels more closely the HIV-1 paradigm than previously thought.

## Materials and Methods

**Plasmids.** APOBEC3B and -3G cDNAs were a gift from Naveenan Navaratnam (Imperial College, London), and APOBEC3C and -3F were obtained as IMAGE clones from the HGMP Resource Centre. The different APOBEC3 cDNAs were subcloned in the expression vector pcDNA3.1D/V5-His-TOPO (Invitrogen). Plasmid pCayw allows expression of pregenomic HBV RNA under control of the powerful cytomegalovirus

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: HBV, hepatitis B virus.

†To whom correspondence should be addressed. E-mail: [simon@pasteur.fr](mailto:simon@pasteur.fr).

© 2005 by The National Academy of Sciences of the USA

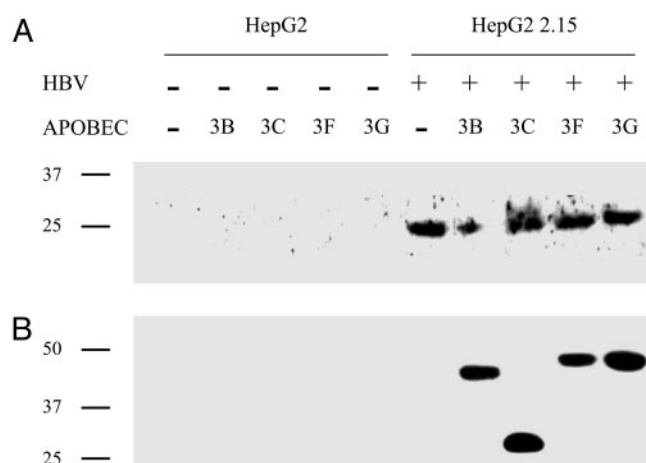
immediate early promoter (CMV-IE) and was provided by Frank Chisari (The Scripps Research Institute) (24). The plasmid pHBVΔcore harbors the same ayw genome and allows HBsAg expression, but is unable to replicate because of a deletion in the core gene (25).

**Cells and Transfections.** Huh7 and HepG2 human hepatoma cell lines were maintained in DMEM complemented with 10% FCS, 50 units/ml penicillin, and 50 μg/ml streptomycin (GIBCO/BRL). HepG2.2.15 cells, a derivative of HepG2 stably expressing HBV (26), were maintained in Williams medium complemented with 10% FCS, 50 units/ml penicillin, 50 μg/ml streptomycin, 0.35 μM hydrocortisone (Sigma), and 5 μg/ml bovine pancreatic insulin (Sigma). Transfections were performed by using FuGENE 6 (Roche). Briefly, 10<sup>5</sup> Huh7 cells were cotransfected with equal amounts (2 μg) of the pCayw and individual plasmids encoding human APOBEC3B, -3C, -3F, or -3G cDNAs. Controls were performed in parallel with pHBVΔcore. Total DNA was recovered 3 days after transfection and extracted.

**Viral Capsid Purification and Western Blotting.** Cells were collected and incubated in PBS with 1% Nonidet P-40 at 4°C overnight (27). After low-speed centrifugation, the lysates were collected and centrifuged at 27,000 × g for 2.5 h. The viral nucleocapsids were resuspended in 60 μl of 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, and 1 mM EDTA. Thirty microliters of purified virus nucleocapsids (≈10<sup>6</sup> cell equivalents) per lane were separated on SDS-polyacrylamide gels and transferred to nitrocellulose (Schleicher & Schuell). As primary antibodies, a mouse monoclonal antibody specific for the HBcAg (provided by Marie-Louise Michel, Institut Pasteur) or a mouse monoclonal antiserum specific for the V5 epitope tag (Invitrogen) were applied overnight after blocking of the membrane in 5% dried milk in PBS. After incubation with the respective horseradish peroxidase-coupled secondary antibody, the membrane was subjected to detection by enhanced chemoluminescence (Pierce).

**PCR Amplification, Cloning, and Sequencing.** A fragment of the HBV genome just 5' to ORF X was amplified by employing a nested procedure. To increase sensitivity and specificity, hot start PCR was performed. The first-round primers were 5'-CG-CAAATATACATCGTATCCAT and 5'-AAGAGTYTYT-TATGTAAGACYTT, where Y is T/C and R is A/G, and primers 5'-ATGGCTGCTARGCTGTGCTGCCAA and 5'-AAGTGCACACGGTYGCGCAGAT were used for the second round. Hypermutated genomes were identified by using differential DNA denaturation PCR (3D-PCR) in a two-round procedure (23). The first reaction involved standard amplification, the reaction parameters were 95°C for 5 min, followed by 35 cycles (95°C for 30 s, 55°C for 30 s, and 72°C for 30 s), and finally for 10 min at 72°C for the first round. Differential amplification occurred in the second round, using the equivalent of 0.5 μl of the first-round reaction as input. Cycling conditions were 90°C or 88°C for 5 min, followed by 35 cycles (90°C or 88°C for 60 s, 45°C for 30 s, and 72°C for 30 s), and finally 10 min at 72°C. The buffer conditions were 2.5 mM MgCl<sub>2</sub>, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 200 μM each dNTP, 100 μM each primer, and 2.5 units of *Taq* DNA polymerase (Applied Biosystems) in a final volume of 50 μl. Second-round PCR yielded a 213-bp fragment.

PCR products were purified from agarose gels (Qiaex II kit, Qiagen) and ligated in the TOPO TA cloning vector (Invitrogen). After transformation of XL1 Blue cells, up to 50 clones were picked. Sequencing was performed by using Thermosequenase (United States Biochemical, Amersham Pharmacia). The 245-bp segment of the ampicillin resistance gene was amplified by using amp1, 5'-ATRCARTCTGCCATAAC-CAT, and amp2, 5'-GGGAAGYTAGAGTAAGTAGTT. The minimum denaturation temperature was 89°C.



**Fig. 1.** APOBEC3 molecules are associated with HBV capsids. (A) Western blot analysis of the purified HBV capsids from HepG2.2.15 transfected by APOBEC3 expression plasmids and detected by anti-HBcAg antibodies. HepG2 cells transfected by APOBEC3 plasmids were used as negative controls. (B) Western blot analysis of the same lysates using a monoclonal antibody to the APOBEC3-tagged V5 epitope. Molecular weight markers (kDa) are shown to the left.

**Patient Material.** Sera were obtained from four chronic carriers. For patient 130.71 the viral load was  $3 \times 10^9$  DNA copies per ml; for patient 12763 the viral load was  $1 \times 10^9$  copies per ml; for patient 129.5 the viral load was  $8 \times 10^8$  copies per ml; and for patient 12741 the viral load was  $6 \times 10^8$  copies per ml. Viral load was determined by the Chiron bDNA method. DNA was phenol extracted and taken up in 10 mM Tris-HCl (pH 8)/0.1 mM EDTA. The equivalent of 7 μl of serum was used for each PCR reaction.

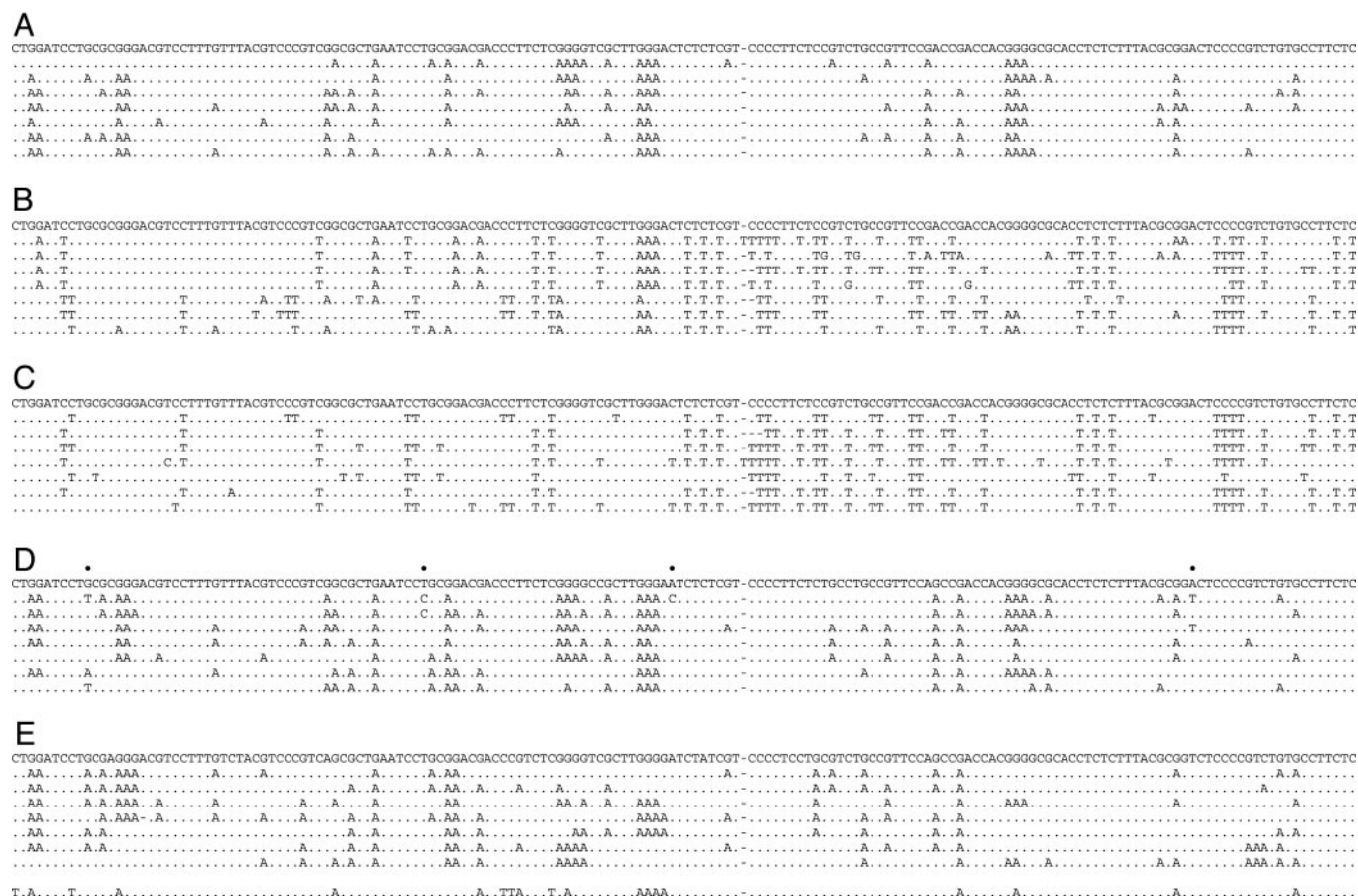
## Results

**APOBEC3 Proteins Are Associated with HBV Capsids.** The cytidine deaminase APOBEC3G has recently been shown to be associated with HBV core antigen (HBcAg) in the cytoplasm of virus-producing cells (20). To ascertain whether the four APOBEC3 enzymes can be associated with the HBV capsid, HepG2.2.15 cells were transfected with expression plasmids encoding V5-tagged forms of APOBEC3B, -3C, -3F, and -3G. Western blot analysis performed 3 days posttransfection confirmed that the purified fractions did indeed contain HBcAg (Fig. 1A). Importantly, all APOBEC3 proteins tested were associated with the purified capsids (Fig. 1B). The specificity of immunodetection is shown by the negative HepG2 controls (Figs. 1A and B).

**Four APOBEC3 Deaminases Can Massively Edit HBV DNA.** Huh7 cells were transfected by the same pCayw HBV plasmid used by Turelli *et al.* (20) along with APOBEC3B, -3C, -3F, and -3G expression plasmids. Given the ability of the 3D-PCR to differentially amplify HIV-1 G → A hypermutants, we reasoned that despite residual plasmid DNA from transfection, the method should be able to selectively amplify *de novo* synthesized and hyperdeaminated HBV DNA from a total DNA extraction. As hyperdeaminated DNA contains multiple uracil residues, *Taq* polymerase was used, as opposed to *Pfu* (20), because the latter has a read-ahead function and stalls at uracil residues in template DNA (28, 29).

Total DNA was recovered from Huh7 cells 3 days after cotransfection. By slowly decreasing the denaturation temperature (Td), it was shown that 93°C was the minimum temperature allowing amplification of the fragment from the pCayw transfection alone. By performing PCR below the minimal denatur-





**Fig. 2.** Massive deamination of HBV DNA by APOBEC3 cytidine deaminases. A selection of hypermutated HBV sequences mapping just 5' to ORF X are shown compared with their reference. The sequence is given with respect to the viral plus strand. Only differences are shown. Hyphens denote gaps introduced to maximize sequence identity. All sequences were unique, indicating that they corresponded to distinct molecular events. (A) Selection of G → A hypermutations resulting from APOBEC3G deamination of minus-strand DNA. (B) A selection of mixed G → A and C → T hypermutants from APOBEC3B, -3F, and -3G cotransfections. (C) A selection of uniquely C → T hypermutants from APOBEC3B, -3F, and -3G cotransfections. (D and E) Hypermutated sequences derived from virions in the sera of patients 130.71 and 12763, respectively. The reference sequence given was the major sequence among 10 derived from PCR products amplified at 95°C. Dots above the reference sequences indicate polymorphic sites among the 10 nonhypermutated sequences. For patient 12763 (Fig. 3E), one of the sequences was a mixed G → A + C → T hypermutant.

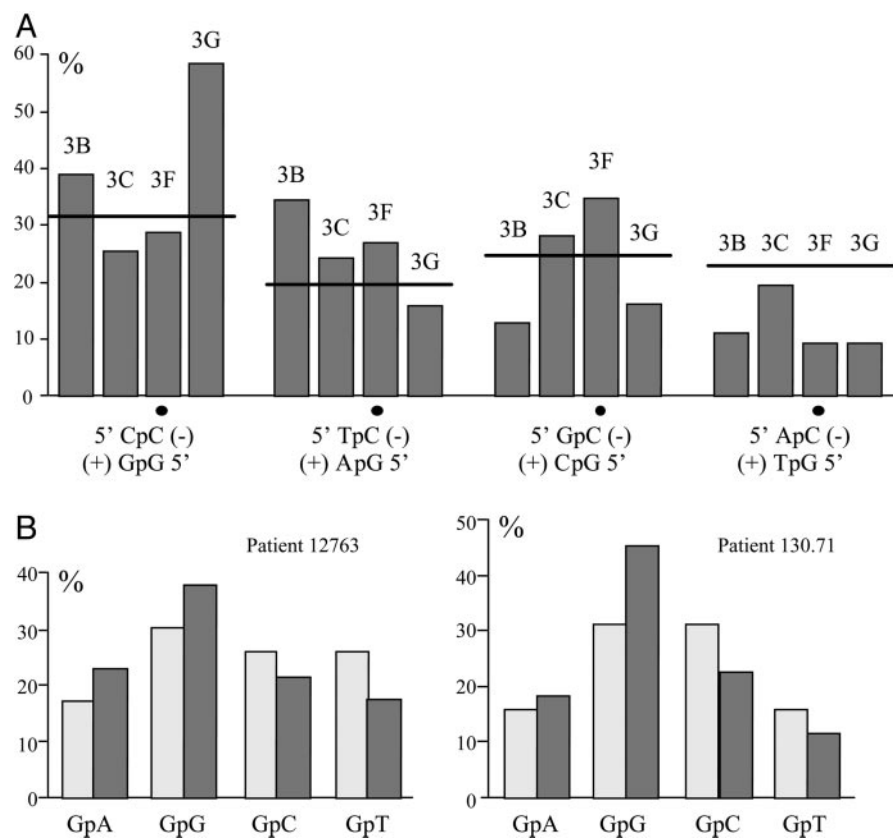
ation temperature, for example 90°C, we were able to amplify DNA from all of the APOBEC3 cotransfections, i.e., APOBEC3B, -3C, -3F, and -3G, but not from pCayw alone. These PCR products were cloned, and DNA from up to 50 clones were sequenced per sample, a small selection being shown for the APOBEC3G sample in Fig. 24. A complete collection of hypermutated sequences is available in Figs. 4–7, which are published as supporting information on the PNAS web site. As expected from the fact that the DNA was amplified at 90°C, all sequences were extensively substituted. All sequences were unique, indicating that they represented independent molecular events. The number of transitions ranged from 14–48 per clone, translating into G → A and/or C → T substitution frequencies of between 13% and 44%.

Because the APOBEC3 genes were overexpressed by way of the powerful CMV-IE promoter, it might be argued that a very small proportion of APOBEC3 molecules found their way to the nucleus. If so, single-stranded DNA in a transcription bubble might be vulnerable to deamination. Given the sensitivity of 3D-PCR, it was possible that it was picking up such rare events. To this end, Huh7 cells were cotransfected by pHBVΔcore and APOBEC3 plasmids. The pHBVΔcore plasmid allows transcription through the locus amplified by 3D-PCR but, due to a deletion in the core gene, was not replication-competent. Total

DNA was harvested at 3 days posttransfection and 3D-PCR performed at 90°C. No DNA was recovered from any of the four cotransfections with pHBVΔcore. These negative findings were confirmed when 3D-PCR was applied to a small locus within the ampicillin resistance gene. Together these observations rule out rare deamination of nuclear HBV DNA in a transcription bubble or deamination of single-stranded DNA resulting from shearing during transfection. Hence, all sequences described here arose from reverse transcription of pregenomic HBV transcripts.

To ascertain the frequency of APOBEC3 editing, limiting dilution of the second-round PCR products (i.e., those at 90°C) showed that the signal titrated out at a frequency of  $\approx 10^{-8}$  for the HBV/APOBEC3 cotransfection experiments (data not shown). By contrast limiting dilution of the first-round products (95°C) titrated out to  $\approx 10^{-11}$ . Given that there must be a little residual plasmid DNA in the samples as well as *de novo* replication of HBV from a few singly transfected cells, the proportion of edited HBV genomes is probably  $> 10^{-3}$ .

**C → T Hypermutants.** Among the APOBEC3 edited HBV sequences were some harboring multiple G → A and C → T substitutions, a selection being shown in Fig. 2*B*. Such sequences were observed for APOBEC3B, -3F, and -3G but not APOBEC3C and represented 7–19% of the total (Figs. 4–7).



**Fig. 3.** Dinucleotide context associated with editing. (A) The ordinate represent the substitution frequency as a function of the 5' nucleotide. Dots indicate the deaminated C residue. The expected value is represented by a horizontal bar and corresponds to the base composition of the locus. A  $\chi^2$  analysis showed that the observed frequencies APOBEC3B, -3F, and -3G deviated significantly from the expected values ( $P < 0.001$ ). (B) Dinucleotide context for G  $\rightarrow$  A hypermutants derived from the serum of patients 12763 and 130.71, respectively. The observed frequencies (dark gray) encompass 481 and 505 transitions for the two patients, respectively. Expected frequencies are shown in light gray.

The simplest interpretation is that the HBV plus strand is also vulnerable to deamination. Because the plus strand of this segment harbors more C than G residues (36.6% vs. 27.2%), it was possible that lowering further the Td might allow amplification of C  $\rightarrow$  T hypermutants alone. Accordingly, the same first-round PCR products, amplified at 95°C, were subjected to PCR at 88°C, as opposed to the 90°C used previously. Again, 3D-PCR products were recovered for the APOBEC3B, -3F, and -3G cotransfections but not for APOBEC3C. Upon cloning and sequencing, uniquely C  $\rightarrow$  T hypermutants were identified, a selection being shown in Fig. 2C.

To rule out editing of HBV RNA that would also produce C  $\rightarrow$  T transitions in the reference strand, RNA was purified from a fraction of the same cell pellets and purified by using an RNase-free DNase protocol. cDNA was made, and the same locus was amplified by PCR at both 95°C and 90°C. Ample DNA was recovered from all four APOBEC3 cotransfections and the HBV control, whereas none could be recovered at the lower temperature, ruling out RNA editing. The C  $\rightarrow$  T hypermutants indicate that HBV plus-strand DNA can also be edited without prior editing of the minus strand, which would give rise to the mixed sequences seen in Fig. 2B.

**Dinucleotide Target Context of APOBEC3 Family Members.** With a large number of hypermutated sequences, it was possible to compare the relative sequence context associated with APOBEC3B, -3C, -3F, and -3G cytosine deamination. As can be seen from Fig. 3A, all four molecules showed subtle differences. The most pronounced effect was for APOBEC3G, where there

was a strong preference for CpC (target is italic). APOBEC3B preferred RpC (R is A,G) over YpC (Y is C,T), APOBEC3F slightly preferred TpC and GpC, whereas APOBEC3C showed no significant bias. Except for APOBEC3F these findings are in keeping with what has been described in the literature (7, 30–34). The same target site specificities were also observed among the G  $\rightarrow$  A and C  $\rightarrow$  T hypermutants, indicating that there was no asymmetry during editing of (+)- and (–)-strand DNA (data not shown).

**G  $\rightarrow$  A Hypermutations in HBV-Infected Patients.** Given the success of 3D-PCR at identifying G  $\rightarrow$  A hypermutants, the technique was applied to DNA extracted from the serum of four chronically HBV-infected patients with high viremia ( $0.6\text{--}3 \times 10^9$  DNA copies per ml). Two of the four samples (130.71 and 12763) proved positive when 3D-PCR was performed at a Td = 90°C. The products were cloned and  $\approx 20$  clones sequenced. As can be seen from a small sample from each patient the sequences were extensively G  $\rightarrow$  A hypermutated (Fig. 2D and E; and see Figs. 8 and 9, which are published as supporting information on the PNAS web site). Interestingly, the extent of hypermutation was comparable to those obtained experimentally (Fig. 2A).

The dinucleotide sequence context for both patients was CpC > TpC > GpC  $\approx$  ApC (Fig. 3B), which is reminiscent of APOBEC3G (Fig. 3A), although this does not preclude a melange of APOBEC3 molecules. Close inspection of the sequences from patient 12763 showed a single G  $\rightarrow$  A and C  $\rightarrow$  T hypermutant (Fig. 2E, bottom sequence) similar to those shown





